

第4节 支持向量机实践

——线性决策面实例



4.1 线性最优分类面实例

例子4.1: 现有样本 (X_i, y_i) , $i = 1, 2, 3, 4$, 如下:

$$X_1 = [\mathbf{x}_{11}, \mathbf{x}_{12}]^T = [0, 0]^T, y_1 = +1$$

$$X_2 = [\mathbf{x}_{21}, \mathbf{x}_{22}]^T = [1, 0]^T, y_2 = +1$$

$$X_3 = [\mathbf{x}_{31}, \mathbf{x}_{32}]^T = [2, 0]^T, y_3 = -1$$

$$X_4 = [\mathbf{x}_{41}, \mathbf{x}_{42}]^T = [0, 2]^T, y_4 = -1$$

如图4.1所示, 求这些样本集的线性最优分类面。





图4.1 分成两类的例4.1样本集示意图

1) 用Fisher线性判别分析求解上述问题:

① 计算两类数据的平均值

输出 $y = +1$ 这类的平均值为 $(\frac{1}{2}, 0)$;

输出 $y = -1$ 这类的平均值为 $(1, 1)$ 。

② 计算两类数据的离差矩阵:

$$A = \begin{pmatrix} -1/2 & 0 \\ 1/2 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

两类的离差矩阵分别为：

$$S_1 = A'A = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \end{pmatrix} \quad S_2 = B'B = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

$$S = S_1 + S_2 = \begin{pmatrix} 2.5 & -2 \\ -2 & 2 \end{pmatrix}$$

③ 解线性方程组：

$$S \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \bar{x}_1^0 - \bar{x}_1^1 \\ \bar{x}_2^0 - \bar{x}_2^1 \end{pmatrix} = \begin{pmatrix} 1/2 - 1 \\ 0 - 1 \end{pmatrix}$$

即

$$\begin{cases} 2.5c_1 - 2c_2 = -0.5 \\ -2c_1 + 2c_2 = -1 \end{cases}$$

得判别系数：

$$c_1 = -3 \qquad c_2 = -3.5$$


④ 根据计算结果，得判别函数为：

$$y = -3x_1 - 3.5x_2$$

⑤ 求出判别临界值：

输出 $y = +1$ 类的平均值对应的判别值：

$$\bar{y}_A = -3 \times 0.5 - 3.5 \times 0 = -1.5$$

输出 $y = -1$ 类的平均值对应的判别值为：

$$\bar{y}_B = -3 \times 1 - 3.5 \times 1 = -6.5$$


从而得到Fisher判别函数的临界值为：

$$y_0 = \frac{-1.5 \times 2 - 6.5 \times 2}{2 + 2} = -4$$

⑥ 最后，得到Fisher算法的分界面方程：

$$-3x_1 - 3.5x_2 + 4 = 0$$



2) 用支持向量机最优分界面求解:

① 分界面函数为: $g(x) = W \cdot X_i + b = 0$


根据统计学习理论, 最优分类面的目标函数为:

$$f(W) = \frac{1}{2}(W \cdot W)$$

约束条件为:

$$y_i \cdot (W \cdot X_i + b) - 1 \geq 0$$

② 求带不等式约束的Lagrange函数及其极值:

$$L(W, b, \alpha) = \frac{1}{2}(W \cdot W) - \sum_{i=1}^n \alpha_i \{y_i[(W \cdot X_i) + b] - 1\}$$



将上述Lagrange函数对参数 W, b 取导，令其为0，有：

$$\begin{cases} \frac{\partial L}{\partial W} = W - \sum_{i=1}^4 \alpha_i y_i X_i = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^4 \alpha_i y_i = 0 \end{cases}$$

③ 原始优化问题转化为参数 α 的极值问题：

$$A(\alpha) = \sum_{i=1}^4 \alpha_i - \frac{1}{2} \sum_{i=1}^4 \sum_{j=1}^4 \alpha_i \alpha_j y_i y_j X_i X_j$$

约束条件为：

$$\sum_{i=1}^4 \alpha_i y_i = 0, \quad \alpha_i \geq 0, \quad i = 1, 2, 3, 4$$


④ 把 X_i 、 y_i 的值代入 $A(\alpha)$ 的表达式，后半部分为：

$$\sum_{i=1}^4 \sum_{j=1}^4 \alpha_i \alpha_j y_i y_j X_i X_j$$

$$= (\alpha_1 \cdot 1 \cdot X_1 + \alpha_2 \cdot 1 \cdot X_2 + \alpha_3 \cdot (-1) \cdot X_3 + \alpha_4 \cdot (-1) \cdot X_4)^2$$

$$= (\alpha_2 X_2 - \alpha_3 X_3 - \alpha_4 X_4)^2$$

$$= \alpha_2^2 X_2^2 - \alpha_2 \alpha_3 X_2 X_3 - \alpha_2 \alpha_4 X_2 X_4 - \alpha_2 \alpha_3 X_2 X_3$$

$$+ \alpha_3^2 X_3^2 + \alpha_3 \alpha_4 X_3 X_4 - \alpha_2 \alpha_4 X_2 X_4 + \alpha_3 \alpha_4 X_3 X_4 + \alpha_4^2 X_4^2$$

$$= \alpha_2^2 X_2^2 - 2\alpha_2 \alpha_3 X_2 X_3 - 2\alpha_2 \alpha_4 X_2 X_4 + \alpha_3^2 X_3^2 + 2\alpha_3 \alpha_4 X_3 X_4 + \alpha_4^2 X_4^2$$



$$\begin{aligned} &= \alpha_2^2 \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 2\alpha_2\alpha_3 \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} - 2\alpha_2\alpha_4 \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ &+ \alpha_3^2 \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} + 2\alpha_3\alpha_4 \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \alpha_4^2 \begin{bmatrix} 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ &= \alpha_2^2 - 4\alpha_2\alpha_3 + 4\alpha_3^2 + 4\alpha_4^2 \end{aligned}$$

从而有：

$$A(\alpha) = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) - \frac{1}{2}(\alpha_2^2 - 4\alpha_2\alpha_3 + 4\alpha_3^2 + 4\alpha_4^2)$$

⑤ 对 $A(\alpha)$ 求偏导, 得到:

$$\left\{ \begin{array}{l} \frac{\partial A(\alpha)}{\partial \alpha_1} = 1 \\ \frac{\partial A(\alpha)}{\partial \alpha_2} = 1 - \alpha_2 + 2\alpha_3 \\ \frac{\partial A(\alpha)}{\partial \alpha_3} = 1 + 2\alpha_2 - 4\alpha_3 \\ \frac{\partial A(\alpha)}{\partial \alpha_4} = 1 - 4\alpha_4 \end{array} \right.$$

令上述偏导为0，并考虑约束条件有：

$$\begin{cases} \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0 \\ \alpha_2 - 2\alpha_3 = 1 \\ -2\alpha_2 + 4\alpha_3 = 1 \\ 4\alpha_4 = 1 \end{cases}$$

从而有：

$$\begin{cases} \alpha_1 + \alpha_2 - \alpha_3 = \frac{1}{4} \end{cases} \quad (1)$$

$$\begin{cases} \alpha_2 - 2\alpha_3 = 1 \end{cases} \quad (2)$$

$$\begin{cases} -2\alpha_2 + 4\alpha_3 = 1 \end{cases} \quad (3)$$

$$\alpha_4 = \frac{1}{4}$$

上面的第(2)、(3)两个方程矛盾。

上面方程组可转化为两个方程组：

$$\text{(I)} \begin{cases} \alpha_1 + \alpha_2 - \alpha_3 = \frac{1}{4} \\ \alpha_2 - 2\alpha_3 = 1 \end{cases} \quad \text{(II)} \begin{cases} \alpha_1 + \alpha_2 - \alpha_3 = \frac{1}{4} \\ -2\alpha_2 + 4\alpha_3 = 1 \end{cases}$$

(I)方程组是一个三元一次方程组，参数只有两个，故有多个解。令 $\alpha_1 = 0$ ，则有：

$$\begin{cases} \alpha_2 - \alpha_3 = \frac{1}{4} \\ \alpha_2 - 2\alpha_3 = 1 \end{cases} \Rightarrow \begin{cases} \alpha_2 - \alpha_3 = \frac{1}{4} \\ \alpha_2 - \alpha_3 = 1 + \alpha_3 \end{cases}$$



求出 $\alpha_3 = -3/4$ 。显然，这个结果和 $\alpha_i \geq 0$ 的函数设置要求矛盾，因此是无效的。

(II) 解方程组(II):

$$\begin{cases} \alpha_1 + \alpha_2 - \alpha_3 = \frac{1}{4} \\ -2\alpha_2 + 4\alpha_3 = 1 \end{cases}$$

在方程组(II)中，令 $\alpha_1 = 0$ ，有：

$$\begin{cases} \alpha_2 - \alpha_3 = \frac{1}{4} \\ \alpha_2 - 2\alpha_3 = \frac{1}{2} \end{cases}$$

从而有： $\alpha_3 = \frac{3}{4}$ ，以及： $\alpha_2 = 1$

最后有：

$$\alpha_1^* = 0, \alpha_2^* = 1, \alpha_3^* = \frac{3}{4}, \alpha_4^* = \frac{1}{4}$$

⑥ 求 W 的最优值 W^* ：

$$\begin{aligned} W^* &= \sum_{i=1}^4 \alpha_i^* y_i X_i \\ &= 0 \cdot 1 \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} + 1 \cdot 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{3}{4} \cdot (-1) \cdot \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \frac{1}{4} \cdot (-1) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -\frac{3}{2} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \end{aligned}$$

⑦ 求参数 b 的最优值 b^*

α_2 、 α_3 、 α_4 的值都大于0，故其对应的样本 X_2 、 X_3 、 X_4 都是支持向量。

从支持向量的约束条件： $y_i \cdot (W \cdot X_i + b) - 1 = 0$

因此，可以从两个不同类别的样本学习得到参数 b 的估计值。

$$\text{当 } y_i = +1 \text{ 类时, 有 } b^* = 1 - W^* \cdot X_i \quad (4)$$

$$\text{当 } y_j = -1 \text{ 类时, 有 } b^* = -1 - W^* \cdot X_j \quad (5)$$

合并公式(4)、(5)式，有：

$$b^* = -\frac{1}{2} W^* (X_i + X_j) \quad (6)$$

其中： X_i 是 +1 类的支持向量， X_j 是 -1 类的支持向量。

假设利用 X_2 、 X_3 样本点来求取 b^* ，则有：

$$\begin{aligned} b^* &= -\frac{1}{2}W^*(X_2 + X_3) \\ &= -\frac{1}{2}W^*\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) = -\frac{1}{2}\begin{bmatrix} -\frac{1}{2} & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \frac{3}{4} \end{aligned}$$

⑧ 求最优分类面：

$$\begin{aligned} d(X) &= W^*X + b^* \\ &= \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{3}{4} = -\frac{1}{2}x_1 - \frac{1}{2}x_2 + \frac{3}{4} = 0 \end{aligned}$$

从而有： $d(x) = 3 - 2x_1 - 2x_2 = 0$ ，该结果如图4.2所示。



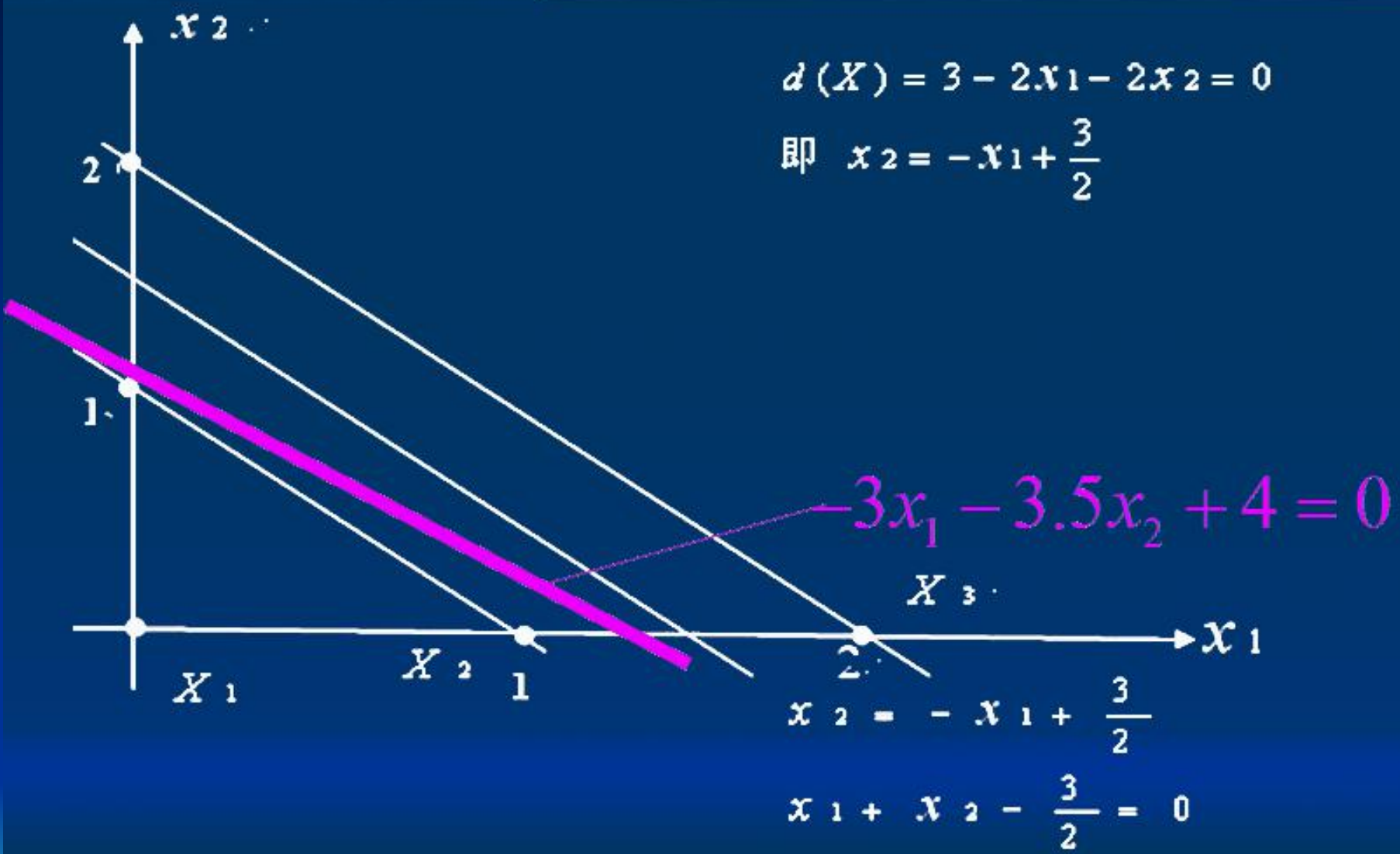


图4.2 Fisher线性分类器和最优分界面示意图

- Fisher准则法得到的结果与最佳分界面求解得到的结果比较，可以发现：
- Fisher准则得到的结果是“**次优**”，而非“**最优**”。这与Fisher算法的“小样本问题”和“次优性问题”有关。



例4.2 如果添加一个数据点，如下所示：

$$X_1 = [x_{11}, x_{12}]^T = [0, 0]^T, \quad y_1 = +1$$

$$X_2 = [x_{21}, x_{22}]^T = [1, 0]^T, \quad y_2 = +1$$

$$X_3 = [x_{31}, x_{32}]^T = [0, 1]^T, \quad y_3 = +1$$

$$X_4 = [x_{41}, x_{42}]^T = [2, 0]^T, \quad y_4 = -1$$

$$X_5 = [x_{51}, x_{52}]^T = [0, 2]^T, \quad y_5 = -1$$

数据分布如图4.3所示，用Fisher法求解线性分类面。



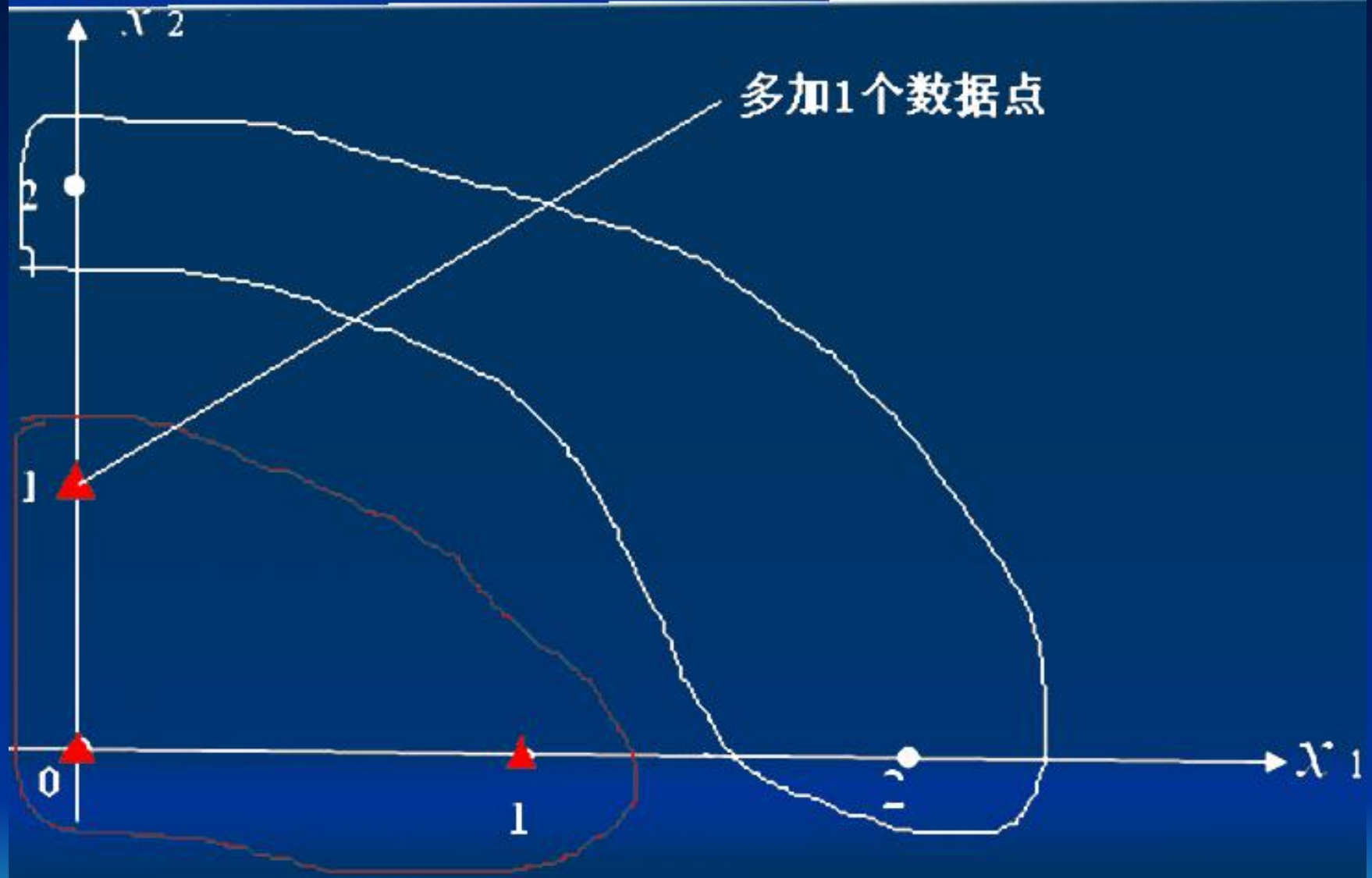


图4.3 添加一个数据点后的数据示意图

1) 求出类内平均值:

输出 $y = +1$ 类的平均值为 $(1/3, 1/3)$;

输出 $y = -1$ 类的平均值为 $(1, 1)$ 。

2) 求解两类数据的离差矩阵:

$$A = \begin{pmatrix} -1/3 & -1/3 \\ 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

两类的离差矩阵分别为：

$$S_1 = A'A = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix} \quad S_2 = B'B = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

$$S = S_1 + S_2 = \begin{pmatrix} 2\frac{2}{3} & -2\frac{1}{3} \\ -2\frac{1}{3} & 2\frac{2}{3} \end{pmatrix}$$

3) 求解Fisher线性方程组:

$$S \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \bar{x}_1^0 - \bar{x}_1^1 \\ \bar{x}_2^0 - \bar{x}_2^1 \end{pmatrix} = \begin{pmatrix} 1/3 - 1 \\ 1/3 - 1 \end{pmatrix}$$

即

$$\begin{cases} \frac{8}{3}c_1 - \frac{7}{3}c_2 = -\frac{2}{3} \\ -\frac{7}{3}c_1 + \frac{8}{3}c_2 = -\frac{2}{3} \end{cases}$$

得判别系数: $c_1 = c_2 = -2$



4) 得判别函数为:

$$y = -2x_1 - 2x_2$$

5) 判别临界值求解:

输出 $y = +1$ 类的平均值对应的判别值:

$$\bar{y}_A = \frac{1}{3} \times (-2) + \frac{1}{3} \times (-2) = -\frac{4}{3}$$

输出 $y = -1$ 类的平均值对应的判别值:

$$\bar{y}_B = -2 \times 1 - 2 \times 1 = -4$$



从而Fisher判别函数的临界值为：

$$y_0 = \frac{-\frac{4}{3} \times 2 - 4 \times 2}{2 + 2} = -\frac{8}{3}$$

6) 最后，得到Fisher算法的分界面方程：

$$-2x_1 - 2x_2 + \frac{8}{3} = 0 \quad \text{即} : -x_1 - x_2 + \frac{4}{3} = 0$$

该结果与最优分类面的比较如图4.4所示。



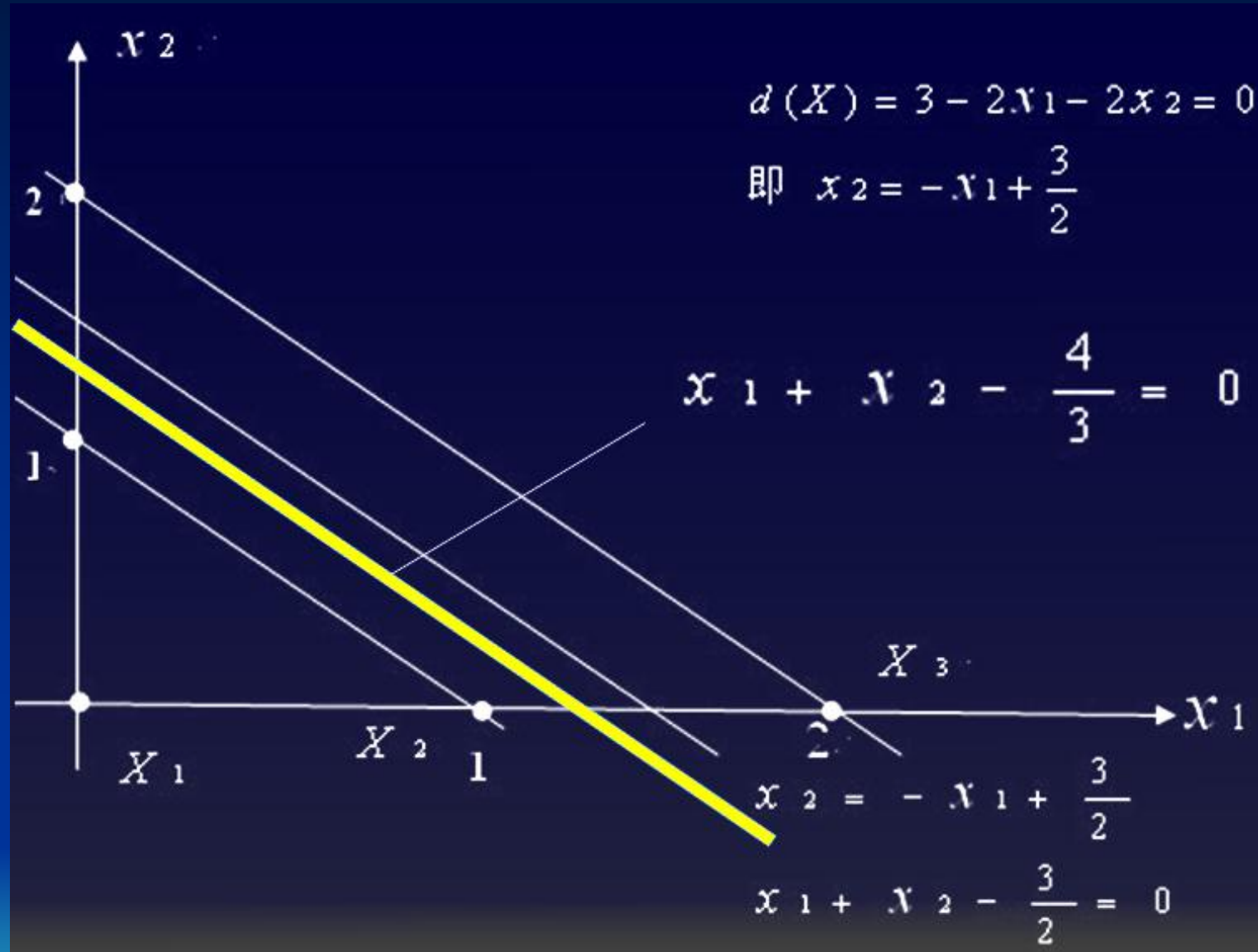


图4.4 添加一个数据点后的Fisher分类面示意图

例4.3: 现有样本 (X_i, y_i) , $i = 1, 2, 3, 4$ 如下:

$$X_1 = [0, 0]^T, y_1 = +1; \quad X_3 = [3, 0]^T, y_3 = -1;$$

$$X_2 = [0, 1]^T, y_2 = +1; \quad X_4 = [0, 3]^T, y_4 = -1。$$

求该样本集的SVM最优分类面。

解: 1) 求 $A(\alpha)$ 的函数表达式:

$$A(\alpha) = \sum_{i=1}^4 \alpha_i - \frac{1}{2} \left(\sum_{i=1}^4 \alpha_i y_i X_i \right)^2$$

其中:

$$\left(\sum_{i=1}^4 \alpha_i y_i X_i \right)^2 = \left(\alpha_1 \cdot 1 \cdot X_1 + \alpha_2 \cdot 1 \cdot X_2 + \alpha_3 \cdot (-1) \cdot X_3 + \alpha_4 \cdot (-1) \cdot X_4 \right)^2$$

$$\begin{aligned}
&= (\alpha_1 X_1 + \alpha_2 X_2 - \alpha_3 X_3 - \alpha_4 X_4)^2 \\
&= (\alpha_2 X_2 - \alpha_3 X_3 - \alpha_4 X_4)^2 \\
&= \alpha_2^2 X_2^2 - 2\alpha_2 \alpha_3 X_2 X_3 - 2\alpha_2 \alpha_4 X_2 X_4 + \alpha_3^2 X_3^2 + 2\alpha_3 \alpha_4 X_3 X_4 + \alpha_4^2 X_4^2 \\
&= \alpha_2^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} - 2\alpha_2 \alpha_3 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} - 2\alpha_2 \alpha_4 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} \\
&\quad + \alpha_3^2 \begin{bmatrix} 3 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} + 2\alpha_3 \alpha_4 \begin{bmatrix} 3 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \alpha_4^2 \begin{bmatrix} 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} \\
&= \alpha_2^2 - 6\alpha_2 \alpha_4 + 9\alpha_3^2 + 9\alpha_4^2
\end{aligned}$$



2) 对 $A(\alpha)$ 求导, 并令其为0, 得到:

$$\left\{ \begin{array}{l} \frac{\partial A(\alpha)}{\partial \alpha_1} = 1 \\ \frac{\partial A(\alpha)}{\partial \alpha_2} = 1 - \frac{1}{2}(2\alpha_2 - 6\alpha_4) = 0 \\ \frac{\partial A(\alpha)}{\partial \alpha_3} = 1 - \frac{1}{2}(18\alpha_3) = 0 \\ \frac{\partial A(\alpha)}{\partial \alpha_4} = 1 - \frac{1}{2}(-6\alpha_2 + 18\alpha_4) = 0 \end{array} \right.$$

3) 求解 α :

$$\begin{cases} \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0 & (1) \\ 1 - \alpha_2 + 3\alpha_4 = 0 & (2) \\ 1 - 9\alpha_3 = 0 & (3) \\ 1 + 3\alpha_2 - 9\alpha_4 = 0 & (4) \end{cases}$$

由公式 (3) 得到: $\alpha_3 = \frac{1}{9}$

上面方程组中(2)、(4)是矛盾方程, 分解成2个方程组:

$$(I) \begin{cases} \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0 \\ 1 - \alpha_2 + 3\alpha_4 = 0 \\ 1 - 9\alpha_3 = 0 \end{cases}$$

$$(II) \begin{cases} \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0 \\ 1 + 3\alpha_2 - 9\alpha_4 = 0 \\ 1 - 9\alpha_3 = 0 \end{cases}$$

解方程组(I):

$$\begin{cases} \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0 \\ \alpha_2 - 3\alpha_4 = 1 \\ 1 - 9\alpha_3 = 0 \end{cases}$$

3 个方程求解4个未知数，因此会有不止一组解。

取 $\alpha_1 = 0$ ，解出： $\alpha_1 = 0, \alpha_2 = -\frac{3}{9}, \alpha_3 = \frac{1}{9}, \alpha_4 = -\frac{4}{9}$

其中： $\alpha_2 < 0, \alpha_4 < 0$

不满足 $\alpha_i \geq 0$ 的要求，故解无效。



解方程组(II):

$$\begin{cases} \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0 \\ 1 + 3\alpha_2 - 9\alpha_4 = 0 \\ 1 - 9\alpha_3 = 0 \end{cases}$$

同方程组(I), 方程组(II)也是3个方程求解4个未知数, 因此会有不止一组解。取 $\alpha_1 = 0$, 解出:

$$\alpha_1^* = 0, \alpha_2^* = \frac{1}{3}, \alpha_3^* = \frac{1}{9}, \alpha_4^* = \frac{2}{9}$$

满足 $\alpha_i \geq 0$ 的要求, 故为有效解。



4) 求最优权向量 W^* :

$$\begin{aligned} W^* &= \sum_{i=1}^4 \alpha_i^* y_i X_i \\ &= \frac{1}{3} \cdot (+1) \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \frac{1}{9} \cdot (-1) \cdot \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \frac{2}{9} \cdot (-1) \cdot \begin{bmatrix} 0 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ +\frac{1}{3} \end{bmatrix} + \begin{bmatrix} -\frac{1}{3} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} \end{aligned}$$

5) 求最优截距 b^* :

$$\begin{aligned} b^* &= -\frac{1}{2}W^*(X_i + X_j) \\ &= -\frac{1}{2}\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \end{bmatrix}\right) \\ &= -\frac{1}{2}\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{2}{3} \end{aligned}$$

6) 求最优分类面方程:

$$\begin{aligned}d(X) &= W^* \cdot X + b^* \\&= \begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{2}{3} \\&= -\frac{1}{3}x_1 - \frac{1}{3}x_2 + \frac{2}{3} \\&= -x_1 - x_2 + 2\end{aligned}$$

即: $d(X) = x_1 + x_2 - 2 = 0$

最优分界面结果如图4.5所示。



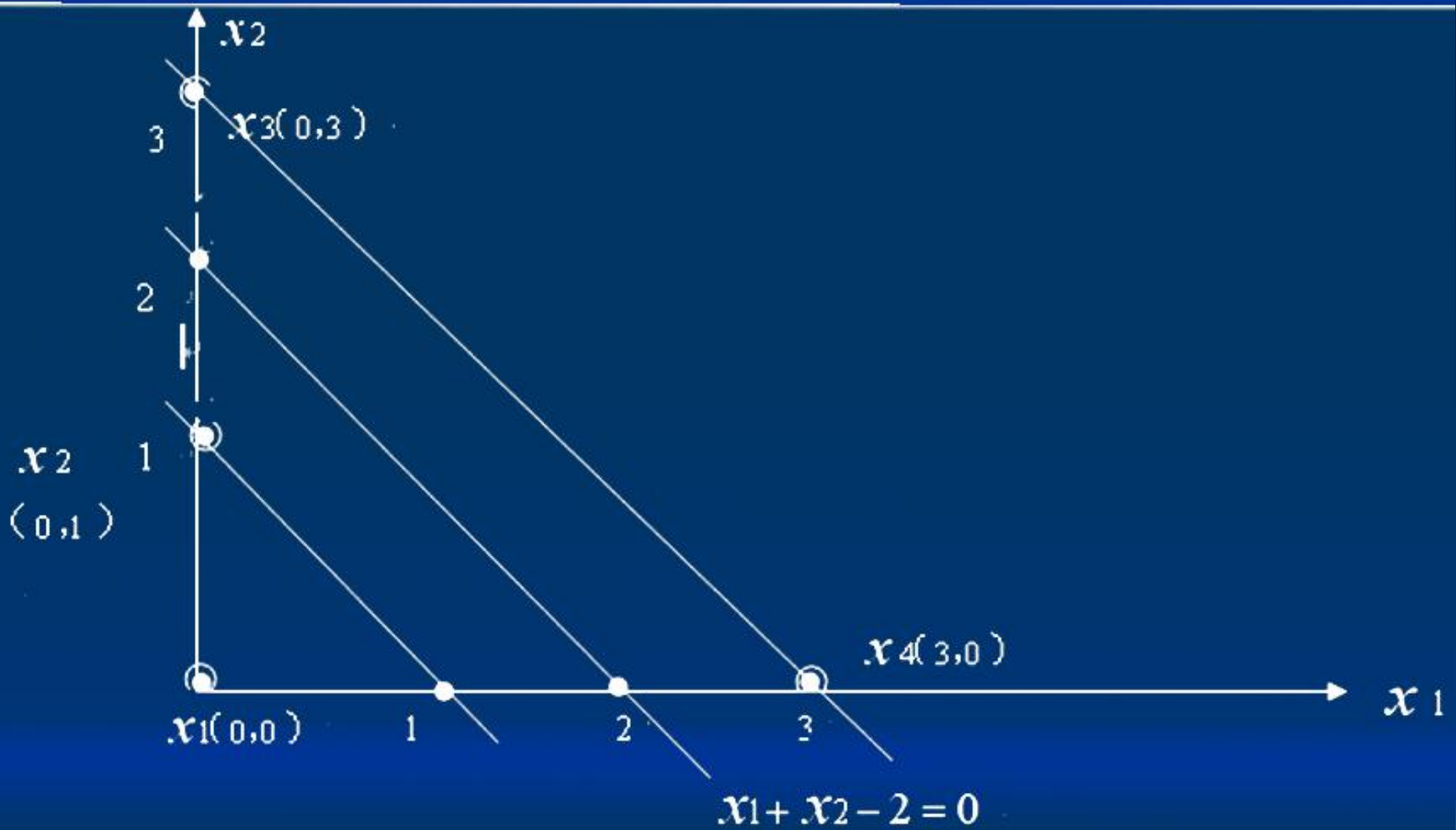


图4.5 例4.3最优分类面结果示意图

4.2 SVM的matlab工具

4.2.1 matlab高级版本中自带的SVM函数

matlab版本为matlab 7.6.0 (R2008a) 这个版本中已经自带SVM算法，分别在生物信息工具箱 (bioinformatics toolbox) 中 `svmclassify` 函数和 `svmtrain` 函数，为前后顺序关系。



- **SVMStruct= svmtrain (Training, Group)**

**% svmtrain的输入为样本点 training 和样本
的分类标签 group ， 输出为一个分类器
svmstruct。**



•Training:

训练集矩阵，每一行对应一个样本（观察量），每一列对应一个特征（属性、变量）。

•Group:

训练集标签（二分类问题标签），可以是一个数值列向量、字符向量或者元胞字符串向量，将训练集中的数据分成两类，标签数目需要与Training中的行数相同，即标签数目需要与样本数目相同，Group中的每一个元素指定了在Training相应行数所代表的样本所属的类别。



- 核函数、核参数、和计算方法等都是可选的，如：

```
SVMStruct = svmtrain(..., 'Kernel_Function', Kernel_FunctionV  
alue, ...)
```

核函数选项如下：

'linear'--默认选项，线性核函数。

'quadratic'--二次核函数。

'rbf'--高斯径向基核函数，默认的scaling因子 $\sigma = 1$ 。

'polynomial'--多项式核函数，默认次数为3。

'mlp'--多层感知机核函数，默认scale和bias参数[1,-1]。



- **SVMStruct**—— 是一个结构体包含训练得到的支持向量机分类器的相关信息，有以下一些域 (fields):

SupportVectors

Alpha

Bias

KernelFunction

KernelFunctionArgs

GroupNames

SupportVectorIndices

ScaleData

FigureHandles



- 然后，将分类器和测试样本testing sample带入svmclassify中，可以得到分类结果和准确度：

```
classes = svmclassify (svmStruct ,  
data(test,:), 'showplot', true);
```

```
% 输出检测样本点的分类结果，且画图表示。
```



代码实例：

- %载入测试数据， Fisher's iris数据

```
load fisheriris      %提取训练集和测试集
data = [meas(:,1), meas(:,2)];
groups = ismember(species,'setosa');
[train, test] = crossvalind('holdOut',groups);
cp = classperf(groups);
```



4.2.2 代码实例:

1) 利用 `svmtrain` 建立支持向量机分类器

```
svmStruct =  
svmtrain(data(train,:),groups(train),'showplot',true)  
title(sprintf('Kernel Function: %s',...  
func2str(svmStruct.KernelFunction)),...  
'interpreter','none');
```

运行结果如图4.6所示。



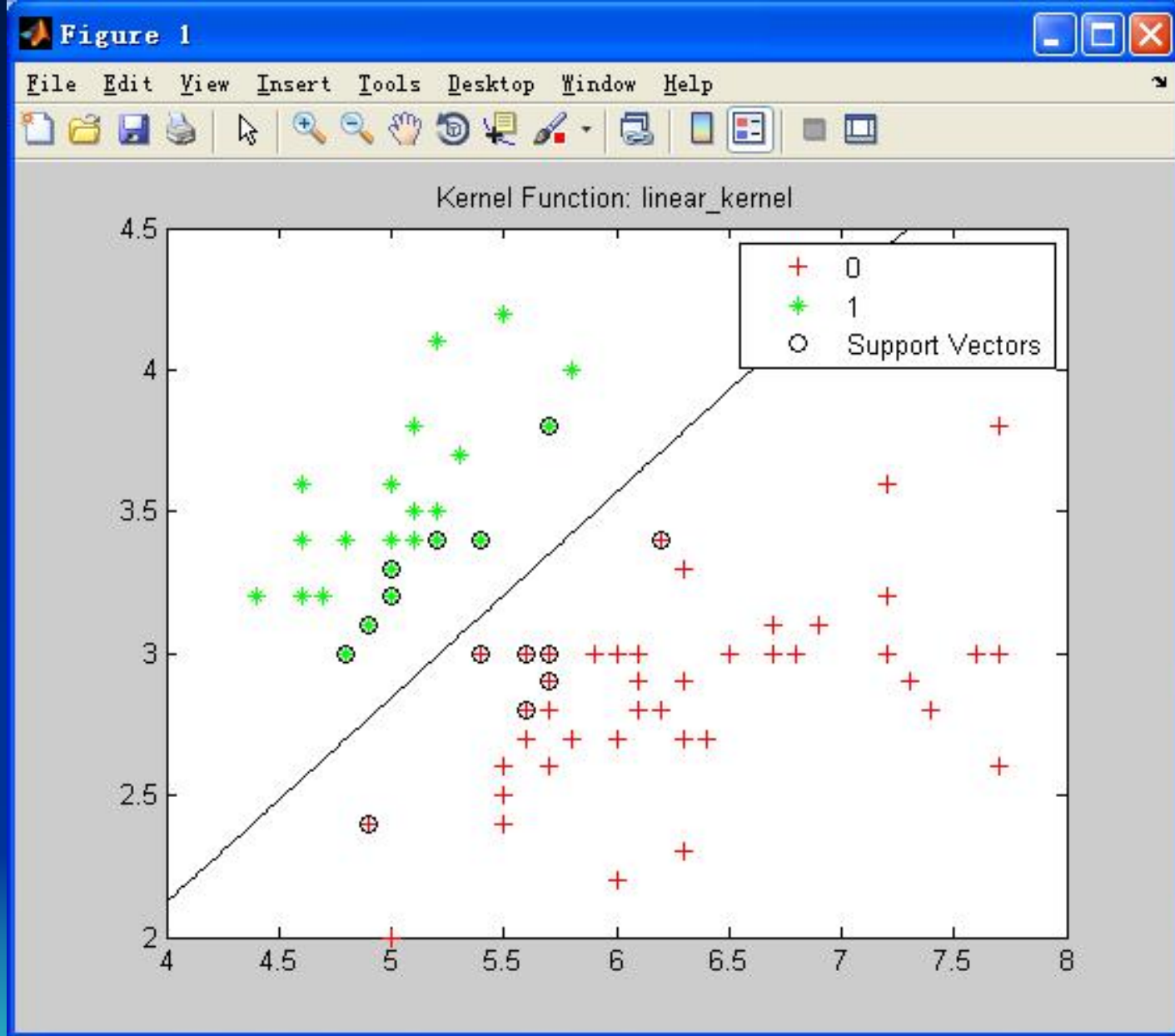


图4.6 matlab的svmtrain函数运行结果示意图

2) %利用得到的svmStruct来进行预测

```
classes =
```

```
svmclassify(svmStruct,data(test,:), 'showplot',true);
```

```
classperf(cp,classes,test);
```

```
cp.CorrectRate
```

```
ans =
```

```
0.9867
```



习题:

- 有两类样本集:

$$X_1^1 = [0 \ 0 \ 0]^T, \quad X_1^2 = [1 \ 0 \ 0]^T, \quad X_1^3 = [1 \ 0 \ 1]^T;$$

$$X_{-1}^1 = [0 \ 0 \ 1]^T, \quad X_{-1}^2 = [0 \ 1 \ 0]^T, \quad X_{-1}^3 = [0 \ 1 \ 1]^T.$$

试分别用Fisher线性判别方法和SVM最优界面法给出这两类样本的分类面。

